

2009

Puntos Técnicos para la Validación y Verificación de una Nueva Institución

Biblioteca Digital Colombiana
Control



Autor:
Laureano Felipe Gómez
2009-06-25



En el desarrollo de las actividades y elaboración de este documento colaboró:

Edwin Montoya – Universidad EAFIT

El proyecto “Creación de la Biblioteca Digital Colombiana” fue financiado por el Ministerio de Educación Nacional, el Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología, y la Red Nacional Académica de Tecnología Avanzada; bajo el contrato I.F. 009-07, de la convocatoria 393 de 2007.

Las universidades participantes en el proyecto son: Universidad ICESI, Universidad EAFIT, Universidad del Norte, Universidad de La Salle, Universidad de Medellín, Universidad de Antioquia, Universidad de La Sabana, Universidad Santiago de Cali, Universidad Autónoma del Caribe, Universidad Nacional de Colombia, Universidad Autónoma de Occidente, Corporación Universitaria Minuto de Dios y Universidad Colegio Mayor de Nuestra Señora del Rosario.



Esta obra está bajo una [licencia de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Si usted tiene algún comentario, sobre este proyecto o alguno de los documentos relacionados, por favor enviar un e-mail a emontoya@eafit.edu.co.

CONTENIDO

INTRODUCCIÓN.....	4
1 ÁREA DE IDENTIFICACIÓN DE LA INSTITUCIÓN.....	5
2 ÁREA DE IDENTIFICACIÓN DE LA FUENTE DE INFORMACIÓN DIGITAL.....	6
2.1 RECURSOS DE LA FUENTE DE INFORMACIÓN.....	6
2.2 VISIBILIDAD WEB.....	8
3 IDENTIFICACIÓN DE LA INFRAESTRUCTURA TECNOLÓGICA.....	10
4 ÁREA DE ANÁLISIS DE LA CATALOGACIÓN / CONTROL DE METADATOS	12
5 ÁREA DE IDENTIFICACIÓN DE PROTOCOLOS DE BÚSQUEDA (INTEROPERABILIDAD)	15

INTRODUCCIÓN

Este documento busca establecer clara y detalladamente los lineamientos técnicos mínimos que deben cumplir las distintas fuentes de información digital que posean las instituciones para su ingreso como miembros de BDCOL.

De esta manera se establece un conjunto de criterios y pautas de validación, evaluación y verificación que permitan establecer un marco común para la identificación de los principales recursos y fuentes de información digitales académicas y de investigación, que estén agrupadas en colecciones digitales (Contenidos Estáticos, CMS, DMS, Repositorios Documentales, Bibliotecas Digitales, Bibliotecas Virtuales y Bibliotecas Electrónicas).

Como instrumento de evaluación/verificación se tiene un formulario, que debe ser llenado por cada recurso/fuente de información digital localizada que agrupen una o varias colecciones digitales, esto permite tener una mayor diseminación de los contenidos presentes en cada una de las instituciones analizadas.

La evaluación/verificación está delimitada en cinco áreas: Identificación de la Institución, Identificación del recurso/fuente de información, Identificación de la infraestructura tecnológica, Colecciones y protocolos de interoperabilidad.

1 ÁREA DE IDENTIFICACIÓN DE LA INSTITUCIÓN

Ítem de Evaluación	Descripción	Datos
1.1. Nombre de la Institución	Corresponde con el nombre de la Institución / Comunidad responsable por la publicación, uso y acceso al recurso/fuente de información digital.	
1.1.1 URL	Corresponde con la dirección web de la institución/comunidad	
1.2. Unidad de Información:	Se debe diligenciar el nombre ó detalles de identificación de la Unidad de información responsable	
1.2.1 URL	Corresponde con la dirección web de la Unidad de Información	
1.3. Persona de Contacto General	Se debe indicar los datos de contacto del director de la Unidad de Información ó la persona encargada de responder por el recurso/fuente de información digital	
1.3.1. Nombre		
1.3.2. Cargo		
1.3.3. Correo		
1.3.4. Teléfono		
1.4. Persona de Contacto Técnico	Se debe indicar los datos de contacto de la persona responsable por la infraestructura tecnológica (Hardware y Software).	
1.4.1. Nombre		
1.4.2. Cargo		
1.4.3. Correo		
1.4.4. Teléfono		

2 ÁREA DE IDENTIFICACIÓN DE LA FUENTE DE INFORMACIÓN DIGITAL

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
2.1. Nombre de la fuente de información Digital:	Indica el nombre genérico dado al recurso/fuente de información digital	
2.2. URL	Corresponde con la dirección WEB del recurso/fuente	
2.3. Descripción	Se debe colocar una pequeña descripción del recurso en sí.	
2.4. Datos de Especificación	Determina si el recurso/fuente analizada corresponde con alguna de las siguientes estructuras:	<ul style="list-style-type: none"> • Colección Digital Estática • Sistema de Gestión de Contenidos (CMS) • Sistema de Gestión de Documentos (DMS) • Repositorio Institucional • Repositorio Temático • Repositorio Financiado • Biblioteca Electrónica • Biblioteca Digital • Biblioteca Virtual • Otro (Wiki, Blog, etc.)

2.1 RECURSOS DE LA FUENTE DE INFORMACIÓN

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
2.5 Cantidad de Documentos Digitales	Si es posible se debe estimar la cantidad de Objetos Digitales que se encuentran actualmente (Colocar la Fecha de Análisis)	

	Cantidad de Objetos Digitales	Tamaño Asociado (en MB)	Nivel de Acceso (Abierto, Cerrado, Mixto)	% de crecimiento Mensual
Libros				
Revistas				
Artículos de Revistas				
Boletines / Folletos				
Tesis				
Mapas				
Fotografías				
Videos				
Música				
Software				
Material de Clases				
Informes de Investigaciones				
Documentos Patrimoniales				
Objetos de Aprendizaje				
Otros				

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
2.6. Formatos/Extensiones de Archivo	Indique cuales son las extensiones / formatos de archivos que manejan.	<ul style="list-style-type: none"> • DOC • XLS • PPT • ODT • ODS • ODP • PDF • GIF • JPG • AVI • MOV • MP3 • WAV • EXE Otros: _____

2.2 VISIBILIDAD WEB

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
2.7 Visibilidad Web	Se busca establecer la visibilidad e impacto del recurso/fuente de información digital	
2.7.1. Google	Indica si aparece referenciado en Google, Google Académico	SI/NO
2.7.2 Metabuscadores	Indica su presencia en sistemas de metabúsqueda de libre acceso en Internet	<ul style="list-style-type: none"> • OAIster • ARC • METALIS • AVANO • Callima • BASE • CASSIR • CitebaseSearch • CYCLADES • ColLoib • Clio-i • DL_KARVEST • DP9 • GEO-LEO • IcITE • MeIND • METALIS • Perseus • PEIADI • SAIL • SDL Search Digital Libraries • SCIENTIFIC COMMONS • PUBLIC KNOWLEDGE HARVESTER • Scirus
2.7.3 Directorios Especializados	Busca establecer si el recurso/fuente aparece referenciado en listados / directorios en Internet	<ul style="list-style-type: none"> • OpenDOAR : http://www.opendoar.org/ • Registry of Open Access Repositories (ROAR): http://roar.eprints.org • The University of Illinois OAI-PMH Data Provider Registry: http://gita.grainger.uiuc.edu/registry/ y Service Provider Registry: http://gita.grainger.uiuc.edu/registry/ • Celestial: http://celestial.eprints.org

-
- OAI-PMH
<http://www.openarchives.org/Register/BrowseSites>
 - Listado repositorios institucionales (IFLA Seúl, 2006):
http://www.ifla.org/IV/ifla72/papers/151-Oliver_Swain-en.pdf
 - Directorio instancias instaladas de DSpace:
<http://wiki.dspace.org/index.php/DSpaceInstances>
 - Directorio instancias instaladas de E-prints:
<http://www.eprints.org/software/archives/>
 - Open Archives Initiative OAI Testing Site Repository Explorer:
<http://re.cs.uct.ac.za/>
 - Repository 66:
<http://maps.repository66.org>
-

3 IDENTIFICACIÓN DE LA INFRAESTRUCTURA TECNOLÓGICA

Esta área permite conocer los detalles tecnológicos de infraestructura en Hardware y Software que soportan el recurso/fuente de información digital.

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
3.1. Datos Básicos del Servidor Principal		
3.1.1. Nombre:	Corresponde con el Nombre del Servidor	
3.1.2. Dirección IP - Internet	Se debe indicar la dirección IP de acceso por Internet	
3.1.3. Dirección IP - RENATA	Se debe indicar la dirección IP de acceso por la red RENATA	
3.2. Detalles de Software		
3.2.1. Software Utilizado	Corresponde con el nombre de la Solución utilizada (si lo posee).	<ul style="list-style-type: none"> • Bepress • CDS Invenio • CDSware • CONTENTdm • Desarrollo Propio • DigiTool • DiVA • DoKS • DPubS • Dspace • EDOC • Eprints • ETD-db • Fedora • Fez/Fedora • Greenstone • HAL • HyperJournal

- i-Tor
- MyCoRe
- OJS (Open Journals System)
- Open Repository
- OPUS (Open Publications System)
- PDLIB
- SciX
- Otro: _____

3.2.2 Versión

Indica el número o nombre de la versión utilizada.

4 ÁREA DE ANÁLISIS DE LA CATALOGACIÓN / CONTROL DE METADATOS

Esta área permite conocer en detalle los componentes sintácticos y semánticos que manejan los diferentes esquemas de metadatos asociados a los objetos digitales.

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
4.1. Esquema de metadatos	Se debe especificar el(los) diferentes modelos/estándares de metadatos que utiliza el recurso/fuente de información.	<ul style="list-style-type: none"> • MARC • DUBLIN CORE • EDTMS • ISAD(G) • EAD • METS • MODS • SCORM • LOM • GEM • OTRO: _____
4.3. Modelos de Codificación de Caracteres de texto	Indica el modelo de codificación de los caracteres de texto que se digitan/muestran en los metadatos	<ul style="list-style-type: none"> • ASCII • ISO8850-1 • UTF-8 • UNICODE
4.4. Vocabularios Controlados	Permite establecer los diferentes vocabularios controlados utilizados que permitan desarrollar un modelo semántico común.	<ul style="list-style-type: none"> • DDC: Corresponde con el Sistema de Clasificación Universal DEWEY: http://www.oclc.org/dewey/ • LCC: Son los esquemas de clasificación del conocimiento desarrollados por la Biblioteca del Congreso de los Estados Unidos de Norteamérica: http://lcweb.loc.gov/catdir/cpsolccco/lcco.html • LCSH: Son el conjunto de términos

(vocabulario controlado) desarrollado por la Biblioteca del Congreso de los Estados Unidos de Norteamérica:
<http://purl.org/dc/elements/1.1/subject>

- MESH: Abreviatura usual, en idioma inglés, de "Medical Subject Headings", una lista de encabezamientos más populares en el área de la Medicina:
<http://www.nlm.nih.gov/mesh/meshhome.html>
- NLM: El conjunto de recursos conceptuales (sistema de clasificación) especificados por la Biblioteca Nacional de Medicina:
<http://wwwcf.nlm.nih.gov/class/>
- TGN: El Tesoro de los nombres Geográficos
- UDC: Corresponde con el sistema de clasificación decimal universal:
<http://www.udcc.org/>
- LEMB: Listado de Encabezamiento de Materias para Bibliotecas:
<http://www.lembdigital.com/>
- TAC: Tesoro Ambiental para Colombia:
<http://www.minambiente.gov.co/tesauro/naveg.htm>
- TDC: El tesoro sobre biodiversidad de Colombia del El Sistema de Información sobre Biodiversidad de Colombia –SIB:
<http://www.siac.net.co/sib/tesauros2/WebModuleTesauros/index.jsp>
- TEE: Tesoro Europeo de la Educación:
<http://www.redined.mec.es/consultas.php>
- UNESCOTHES: El Tesoro de la UNESCO es una lista controlada y estructurada de términos para el análisis temático y la búsqueda de documentos y publicaciones en los campos de la educación, cultura, ciencias naturales, ciencias sociales y humanas, comunicación e

		<p>información: http://databases.unesco.org/thessp/index.html</p> <ul style="list-style-type: none"> • DECS: Descriptores en Ciencias de la Salud, corresponde al tesoro desarrollado por Bireme/OPS que está basado en MESH: http://www.bireme.org • APA: Tesoro de la Asociación de Psicólogos Americanos • Otros: _____
4.8. Sistema de Identificación digital	Corresponde con los distintos modelos para la identificación (No necesariamente Acceso Directo) del Objeto Digital y sus documentos asociados	<ul style="list-style-type: none"> • ID Interno • URI (URN / URL) • DOI • HANDLE • ISBN / ISSN • PURL • OpenURL • POI • XRI • ARK • SICI • ISAN • ISMN • ISRC • ISWC • Otros: _____

5 ÁREA DE IDENTIFICACIÓN DE PROTOCOLOS DE BÚSQUEDA (INTEROPERABILIDAD)

En esta área se evalúan los siguientes protocolos: OAI-PMH, Z39.50, SRU/SRW, RSS, OPENSEARCH y Acceso a los objetos digitales vía HTTP/Web Services. Únicamente se deben diligenciar los protocolos que REALMENTE tenga implementados la fuente de información digital.

Ítem de Evaluación	Descripción	Posibles Valores/ Variables
5.1 Protocolo OAI-PMH		
5.1.1. IP/URL de Acceso:	IP/URL del script OAI-PMH	Texto URL
5.1.2. Esquemas de Metadatos Utilizados	Especifica los diferentes conjuntos/colecciones de modelos/estándares de metadatos manejados (Verbo ListMetadataFormats)	<ul style="list-style-type: none"> • OAI_DC • MARCXML • RFC1807 • DUBLIN CORE • EAD • LOM • OTRO: _____
5.1.3. Observaciones Generales	Permite describir otros aspectos no contemplados	
5.2. Protocolo Z39.50		
5.2.1 IP/URL de Acceso:	Dirección IP del servidor Z39.50	
5.2.2. Puerto	Corresponde con el número de puerto/socket en el cual se encuentra instalado el servidor Z39.50	
5.2.3. Versión	Permite especificar la versión del protocolo que maneja la fuente	
5.2.4. Formatos de Metadatos Soportados	Corresponde con los distintos formatos de metadatos aceptados, en los cuales se transfieren los registros bibliográficos	<ul style="list-style-type: none"> • USMARC • IBERMARC • SUTRS • GRS-1 • MARCXML • GILS

		<ul style="list-style-type: none"> DUBLIN CORE OTROS: _____
5.2.5. Nombre de la Base de Datos	Se debe indicar el nombre que identifica a la base de datos/documentos	
5.2.6. Datos de validación (Usuario / Contraseña)	Especifique/indique si el servidor requiere validación de usuario/contraseña	(Usuario / Contraseña)
5.2.7. Codificación de Carácteres	Corresponde con el conjunto de Carácteres en los cuales se encuentran codificados los registros	<ul style="list-style-type: none"> Utf-8 (Unicode) MARC-8 (ALA) ISO 8859-1 (Latin 1)
5.2.8. Observaciones Generales	Permite describir otros aspectos no contemplados	
5.3. Protocolo SRU/SRW		
5.3.1. IP/URL de Acceso:	IP/URL donde se ubica el script SRW/SRU	
5.3.2. Versión	Permite especificar la versión del protocolo que maneja la fuente	
5.3.3. Observaciones Generales	Permite describir otros aspectos no contemplados	
5.4. Protocolo RSS		
5.4.1. IP/URL de Acceso:	IP/URL donde se ubica el script RSS	
5.4.2. Versión	Permite especificar la versión del protocolo que maneja la fuente	<ul style="list-style-type: none"> 0,91 1,0 2,0 Atom
5.4.3. Esquemas de Metadatos	Especifica los diferentes conjuntos/colecciones de modelos/estándares de metadatos manejados.	<ul style="list-style-type: none"> dc marcxml ead Otros: _____
5.4.4. Esquema de Granularidad	Permite establecer cómo se debe solicitar los diferentes registros de metadatos en base a su fecha.	
5.4.5. Esquema de Búsqueda	Corresponde con las opciones/parámetros que permiten delimitar una búsqueda en RSS	
5.4.6. Codificación de Carácteres	Corresponde con el conjunto de Carácteres en los cuales se encuentran codificados los registros	<ul style="list-style-type: none"> UTF-8 UNICODE-16 ISO 8859-1 ASCII Otros: _____
5.4.5. Observaciones Generales	Permite describir otros aspectos no contemplados	
5.5. Protocolo HTTP (Web Services)		
5.5.1 IP/URL de Acceso:	Dirección IP del servidor HTTP	
5.5.2. Puerto	Corresponde con el número de puerto/socket en el cual se encuentra	

	instalado el servidor	
5.5.3. Ruta de los archivos	Corresponde con la ruta lógica en la cual se encuentran los diferentes objetos digitales	
5.5.4. Funciones/ recursos WSDL Asociados	WSDL describe la interfaz pública de las funciones/procedimientos asociados a los servicios Web	
5.5.5. Observaciones Generales	Permite describir otros aspectos no contemplados	
5.7. Protocolo OpenSearch / A9		
5.7.1 IP/URL de Acceso:	Dirección IP del servidor A9	
5.7.2. Archivo especificación del servicio	Describe la Ruta donde se encuentra el archivo XML que especifica cómo funciona el servicio OpenSearch	
5.7.3. Modelo de Sindicación	Permite especificar el modelo de diseminación de los resultados	<ul style="list-style-type: none"> • 0,91 • 1,0 • 2,0 • Atom
5.7.4. Esquemas de Metadatos	Especifica los diferentes conjuntos/colecciones de modelos/estándares de metadatos manejados.	<ul style="list-style-type: none"> • dc • marcxml • ead • Otros: _____
5.7.5. Observaciones Generales	Permite describir otros aspectos no contemplados	